

NBA BETTING

A QUANTSC PROJECT



Jonathan Qin, Zayn Siddique, Rithik Pothuganti,
Roy Hayyat, Claude (Hana) Yoo

TABLE OF CONTENTS

01

MOTIVATION

Why bet on the NBA?

DATA + MODELS

Data collection, features,
techniques & results

03

02

BACKGROUND

Basketball talk and betting
processes.

CONCLUSIONS

Evaluation and looking
ahead.

04



01

MOTIVATION

MOTIVATION

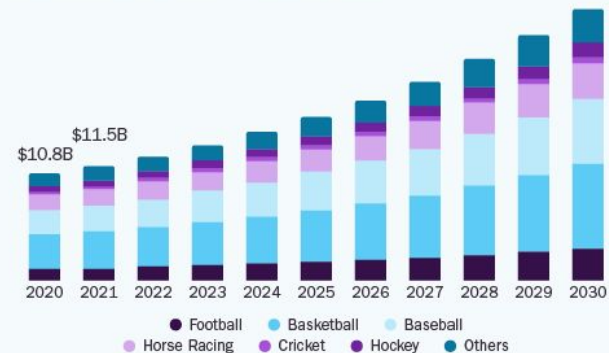
Why sports betting?

- 83.65 billion USD global market in 2022
- 10.3% CAGR until 2030
- Increased online betting
- Largely uninformed bettors
- In the US, basketball has the largest betting market



U.S. Sports Betting Market

size, by sports type, 2020 - 2030 (USD Billion)



MOTIVATION

Why the NBA?

Basketball is the most rapidly evolving sport in the top 5 American leagues. In the last 10 years, the sport has changed more than it ever has, and data has been behind that much of that change. We wanted to see what metrics help teams win the most basketball games.



02

BACKGROUND

BACKGROUND

How does basketball work?

- Basketball is played with 2 teams of 5 players each.
- Each team can score with a 1, 2, or 3 point shot.
- The team with the most points after time is up is the winner.



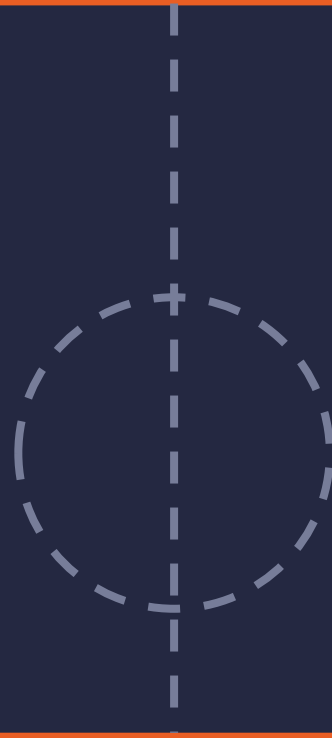
Basketball Terminology

- Points
 - Weighted total of all free throws(1 point), 2-pointers, and 3-pointers
- Rebound
 - Grabbing the ball after a missed shot to start a new possession
- Assist
 - Leading pass that directly leads to a made shot
- Block
 - Deflection of a shot
- Steal
 - Deflection of a pass or taking the ball from the ball-handler resulting in change of possession



Basketball Terminology

- **STAT%**
 - Can be applied to counting stats (AST%, STL%)
 - Measures the percentage of the available stat the player performs
- **FG%**
 - Shots made/Shots taken as a percentage
- **eFG/TS%**
 - Modified FG% to weight shots by value
- **ORTG/DRTG**
 - Points produced/allowed per 100 possessions
- **USG%**
 - Weighted percentage of possessions that end in a player's hands
- **BPM(Box +/-):**
 - Difference between points scored and points given up



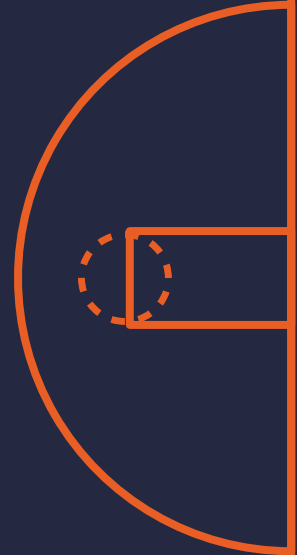
NBA Sports Betting Rundown



Odds and Betting Lines

Odds can be positive and negative.

- Negative odds tell you how much you have to bet **to win** \$100 in profit
- Positive odds tell you how much profit **you win** for every \$100 you bet



Types of Bets

Moneyline

A straight up bet on which team is going to win the game.

GSW -200

OKC +180

Over / Under

Predicting whether the total points scored by both teams combined will be over or under a set total by a sports book.

GSW vs OKC: 210.5

Point Spread

Betting on the margin of points.

GSW +2.5

OKC -2.5

- The underdog can lose by and still cover the spread. Ex: OKC have to not lose by 3 or more points to cover the spread.
- The favorite must win by in order to cover the spread. Ex: GSW have to win by 3 or more points to cover the spread.

TODAY

SPREAD

TOTAL

MONEYLINE

SGP LIVE 0:43 | 3RD QUARTER



ATL Hawks

82

+10.5 -105

O 231.5 -105

+800



BOS Celtics

90

-10.5 -125

U 231.5 -125

-1450

SGP LIVE 8:27 | 1ST QUARTER



MIN Timberwolves

8

+7.5 -105

O 218.5 -110

+280



DEN Nuggets

5

-7.5 -125

U 218.5 -120

-360

SGP 7:00PM



LA Clippers

+12.5 -110

O 225 -110

+575



PHO Suns
















-12.5 -110

U 225 -110

-800

NBA Finals Odds

NBA Championship 2023 - Outright Betting

 Boston Celtics	+195	 Phoenix Suns	+390	 Milwaukee Bucks	+600
 Golden State Warriors	+700	 Philadelphia 76ers	+850	 Denver Nuggets	+950
 Los Angeles Lakers	+1400	 New York Knicks	+3100	 Sacramento Kings	+6000
 Miami Heat	+6000	 Memphis Grizzlies	+10000	 Cleveland Cavaliers	+10000
 Minnesota Timberwolves	+50000	 Atlanta Hawks	+50000	 Los Angeles Clippers	+50000

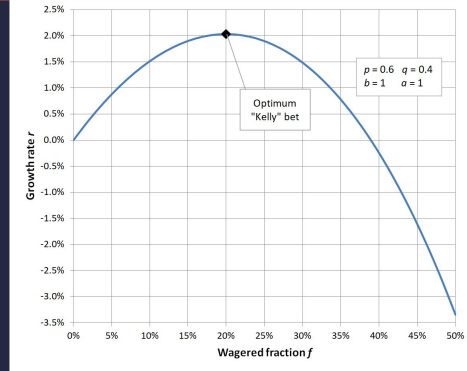
MAY 30, 6:59PM ET

Kelly Criterion

Probability theory - method for sizing up bets, used for determining how much of a predetermined asset to allocate to a bet, ensuring greatest chance of success.

Takes into account the probability of winning/losing a bet, as well as the amount associated.

The criterion suggests that the optimal size of a bet is equal to the percentage advantage f multiplied by the available capital. Here, the decimal odds b represents the amount one wins for every \$1 wagered.



$$f = \frac{bp - q}{b}$$

Where:

f = the fraction of the bankroll to bet

b = the decimal odds

p = the probability of winning

q = the probability of losing, i.e. $1 - p$



03

DATA

Data Collection

- Data collected from Basketball Reference
- Scraped off of the website's tables using the BeautifulSoup library for Python
- Collected regular season/playoff data from 2015-2023
- Collected aggregate season data from 2002-20022
- Data includes
 - Team standings
 - Team stats
 - Player counting/advanced statistics
 - Other miscellaneous data



Models

❖ **Neural Network**

3 layer network

❖ **K Nearest Neighbors**

Similar teams win every year?

❖ **Random Forest Regressor**

Predictions without overfitting



Features

Used NBA data from 2002-2022

Advanced Stats

- TS%, TOV%, FT rate, 3PA rate, Offensive rating, Defensive rating

Conventional Stats/ratios

- Wins, Losses, Free throw/Field Goal ratio, rebound percentage

Miscellaneous metrics

- Attendance/tickets sold, age

K Nearest Neighbors

- Used 100 nearest neighbors
- Teams with similar play-styles tend to win

Metrics

- Mean Squared Error: 11.900
- R-squared: 0.577

Predictions

BOS_2023: 6.67
CLE_2023: 6.53
PHI_2023: 4.47
MEM_2023: 4.74
MIL_2023: 5.48
DEN_2023: 4.62
NYK_2023: 3.05
SAC_2023: 3.13
PHO_2023: 2.34
GSW_2023: 2.01
BRK_2023: 1.66
LAL_2023: 1.48
LAC_2023: 1.82
ATL_2023: 0.85
MIN_2023: 0.99
MIA_2023: 1.19

Neural Network Predictor

Architecture

- 3 Relu Activation layers
- 3 Hidden layers
- 3 Fully connected layers
- 3 Batch normalization layers
- 3 Dropout layers

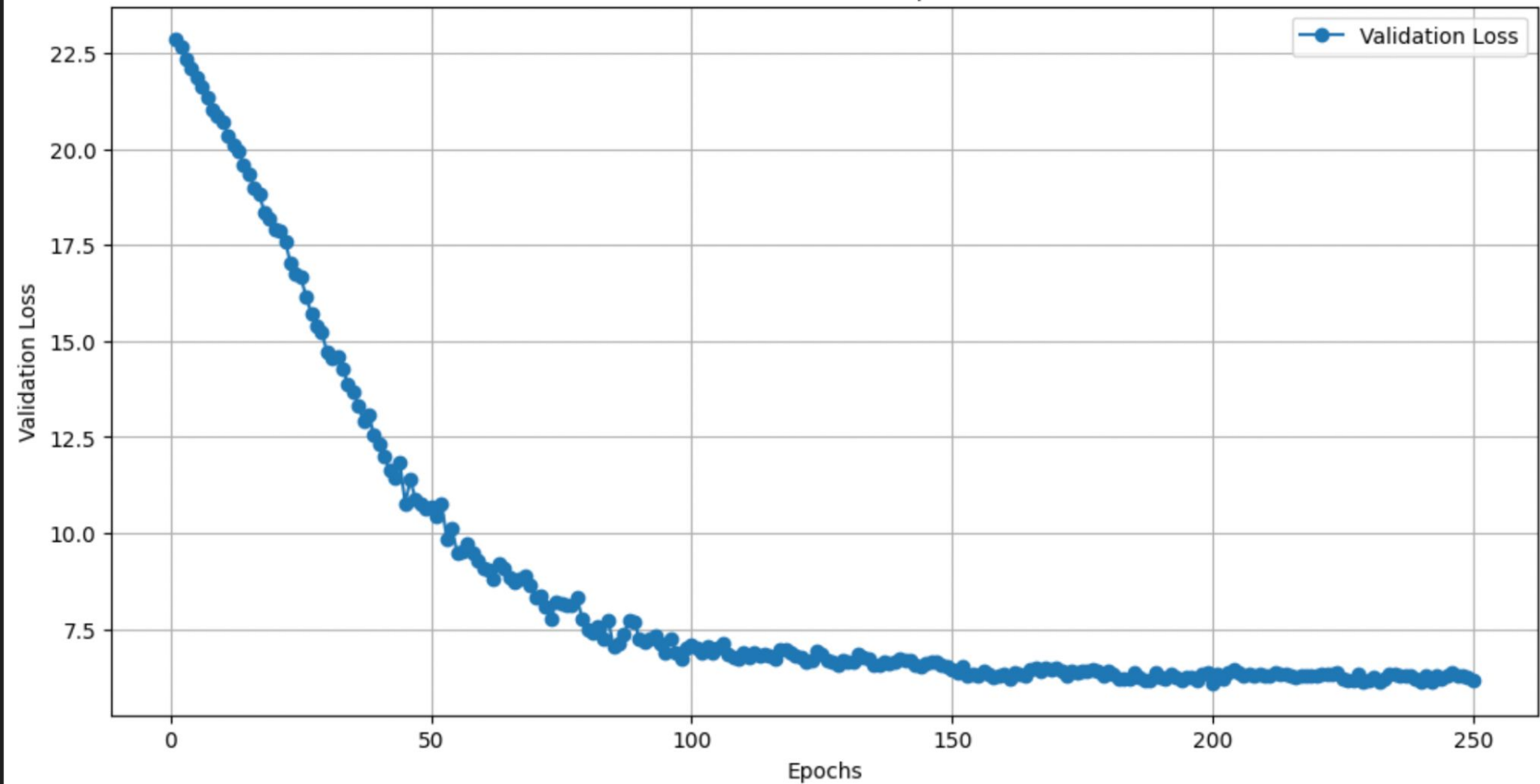
Metrics

- Loss: 6.1753
- R-squared value: 0.6809

Predictions

BOS_2023: 8.318
CLE_2023: 6.412
PHI_2023: 6.488
MEM_2023: 4.865
MIL_2023: 6.685
DEN_2023: 5.819
NYK_2023: 3.034
SAC_2023: 4.410
PHO_2023: 1.680
GSW_2023: 2.267
BRK_2023: 0.884
LAL_2023: 0.718
LAC_2023: 2.173
ATL_2023: -0.184
MIN_2023: -0.080
MIA_2023: 1.733

Validation Loss vs. Epochs



Random Forest Regressor

- Ensemble of decision trees
- Max depth of 10
- 100 trees in forest

Metrics

- Mean Squared Error: 5.270
- R-squared: 0.726

Predictions

BOS_2023: 10.88
CLE_2023: 6.90
PHI_2023: 7.69
MEM_2023: 6.64
MIL_2023: 8.06
DEN_2023: 8.88
NYK_2023: 2.19
SAC_2023: 4.06
PHO_2023: 0.54
GSW_2023: 2.35
BRK_2023: 3.13
LAL_2023: 1.04
LAC_2023: 3.03
ATL_2023: 0.56
MIN_2023: 1.79
MIA_2023: 2.10



04

CONCLUSIONS

WESTERN CONFERENCE



PLAYOFFS

EASTERN CONFERENCE



Nuggets

Suns

Kings

Grizzlies

Nuggets

Nuggets

Celtics

Celtics

Grizzlies

Bucks

Celtics

Bucks

Cavaliers

76ers

Celtics



Drawbacks of models/features

- Does not consider trades
- Has no info about injuries
- Does not factor past playoff performances into predictions
- “Playoff Performers”
- Does not make game-to-game predictions, so less opportunities to place bets

Looking Forward:

- Regular Season betting
 - more predictable
- Live betting



THANK YOU